ABSTRACT

A system and method are provided for processing audio and speech signals using a pitch and voicing dependent spectral estimation algorithm (voicing algorithm) to accurately represent voiced speech, unvoiced speech, and mixed speech in the presence of background noise, and background noise with a single model. The present invention also modifies the synthesis model based on an estimate of the current input signal to improve the perceptual quality of the speech and background noise under a variety of input conditions. The present invention also improves the voicing dependent spectral estimation algorithm robustness by introducing the use of a Multi-Layer Neural Network in the estimation process. The voicing dependent spectral estimation algorithm provides an accurate and robust estimate of the voicing probability under a variety of background noise conditions. This is essential to providing high quality intelligible speech in the presence of background noise.